N94-35630

Measuring Attitude with a Gradiometer

David Sonnabend* Thomas G. Gardner†

Abstract

This paper explores an idea of S. Kant of Goddard — can a gravity gradiometer measure the attitude of a satellite, given that the gravity field is accurately known? Since gradiometers actually measure a combination of the gradient and attitude rate and acceleration terms, the answer is far from obvious. The paper demonstrates yes, and at microradian accuracy. The technique employed is dynamic estimation, based on the momentum biased Euler equations. The satellite is assumed nominally planet pointed, and subject to control, gravity gradient, and partly random drag torques. The attitude estimator is unusual. While the standard method of feeding back measurement residuals is used, the feedback gain matrix isn't derived from Kalman theory. Instead, it's chosen to minimize a measure of the terminal covariance of the error in the estimate. This depends on the gain matrix, and the power spectra of all the process and measurement noises. An integration is required over multiple solutions of Lyapunov equations.

1 Notation & Units

Uppercase bold roman letters are 2 dimensional arrays; e.g., **F**. Lowercase bold roman or greek letters are column vectors; e.g., **r**. Magnitudes of vectors are non–bold; e.g., $r = |\mathbf{r}|$. Lowercase greek subscripts are indices. The Einstein summation convention is used for repeated lower case greek indices. Overdots signify time derivatives; e.g., $\dot{x} = dx/dt$. A T superscript denotes transpose. Primes denote scaled variables. Sines and cosines are denoted by s and c respectively.

$\mathbf{A}$ = constant matrix in Riccati equation
$\mathbf{a}_e = \mathbf{f}_e/m$ = external non–gravitational acceleration on spacecraft
$\mathbf{a}_i$ = inertial acceleration of the ith accelerometer
$\mathbf{B}$ = process noise state distribution matrix
$\mathbf{C}$ = vector of state concern values; C_t = settling time concern value
$\mathbf{D}(\omega)$ = matrix satisfying Lyapunov equation (72)
$E(\mathbf{x})$ = expectation of $\mathbf{x}$; $\mathbf{e}_\alpha^x$ = unit vector along axis α in coordinate system x
$\mathbf{F}$ = plant matrix; $\mathbf{f}_e$ = external non–gravitational force on spacecraft
G = universal gravitational constant = 6.67259×10^{-11} N-m^2/kg^2
$\mathbf{g}$ = gravity field vector at $\mathbf{r}$; $\mathbf{g}_i$ = gravitational acceleration at the ith accelerometer
$\mathbf{H}$ = measurement partials matrix; h = spacecraft pitch momentum bias
$\mathbf{I}_n$ = identity tensor of order n; $\mathbf{J}$ = overall spacecraft inertia tensor
$\mathbf{K}$ = filter feedback gain matrix; $\mathbf{L} = \mathbf{H}^T\mathbf{M}^{-1}\mathbf{H}$ in Riccati equation
k_{1-4} = constants defined in (23); k_f = air drag force constant defined in (33)
$\mathbf{M}$ = combined "equivalent" white noise matrix defined in (80)
m = spacecraft mass; also field source mass
$\mathbf{N}(\omega)$ = matrix defined in (75); $\mathbf{P}_\xi$ = terminal covariance of the error of the estimate
$\mathbf{Q}(\omega)$ = function of power spectra defined in (68); q = performance index; also dynamic pressure
$\mathbf{R}(\tau)$ = autocorrelation matrix with delay τ; $R(0)$ = average power
$\mathbf{r}$ = field position vector relative to m; $\mathbf{r}_{cp}$ = spacecraft center of pressure
$\mathbf{S}(\omega)$ = general noise power spectrum; $\mathbf{S}_v$, $\mathbf{S}_w$ = white noise spectra
s, t superscripts signify spacecraft and trajectory coordinates
t = time in seconds; $t' = t/C_t$ = scaled time; t_s = filter settling time
$\mathbf{U}$ = process noise measurement distribution matrix; $\mathbf{u}$ = vector of controls
$\mathbf{V} = \mathbf{B}\mathbf{S}_w\mathbf{U}^T$ = white process noise effect matrix

*President, Analytical Engineering Co., Boulder, CO, 303-530-9641
†Ph. D. Candidate, University of Colorado

v_o = satellite orbital speed
$\mathbf{W} = \mathbf{B} - \mathbf{KU}$ = process noise effect matrix
$\mathbf{w}$ = process noise vector; w_d = dimensionless air drag random process
$\mathbf{X} = \mathbf{F} - \mathbf{VM}^{-1}\mathbf{H}$ = linear term matrix in Riccati equation
$\mathbf{x}$ = state vector; $\hat{\mathbf{x}}$ = estimate of $\mathbf{x}$; $\dot{x} = dx/dt'$
$\mathbf{Y}$ = measurement noise distribution matrix
$\mathbf{Z} = \mathbf{F} - \mathbf{KH}$ = observer system matrix; $\mathbf{z}$ = vector of measurements
$\boldsymbol{\epsilon}$ = variation in spacecraft $\boldsymbol{\omega}$
$\boldsymbol{\Gamma}$ = gravity gradient tensor; $\Gamma_0 = Gm/r^3$ = gradient scalar due to mass m at distance r
$\lambda(\mathbf{Z})$ = eigenvalue of $\mathbf{Z}$; $\sigma = \Re(\lambda)$ = real part of eigenvalue
$\mu_e = Gm_e$ = gravitational constant of the earth = 3.98603×10^{14} m^3/s^2
$\boldsymbol{\xi} = \hat{\mathbf{x}} - \mathbf{x}$ = error in the state estimate; $\boldsymbol{\tau}_e$ = non-gravitational external torque
$\boldsymbol{\omega}$ = spacecraft angular velocity; ω = angular frequency used in power spectra
ω_c = break frequency in power spectrum; ω_o = orbital mean motion

Unless otherwise stated, the units used in this paper are SI. However, we have also followed common practice in the field of gradiometry on the units of gravity gradient. The natural SI unit is (m/s^2)/m, or just s^{-2}. Since gradient components at the earth's surface are on the order of 1.5×10^{-6} s^{-2}, and are routinely measured to 10^{-9} s^{-2}, or better, this has proved unwieldy. There has now been world wide acceptance of the Eötvös unit: 1 E = 10^{-9} s^{-2}. Here, the SI unit will be used everywhere in the formulas; but Eötvös units will be generally employed in the text.

2 Static Attitude Estimation

The gravitational potential due to a particle of mass m at a distance r is:

$$\Phi = -Gm/r \tag{1}$$

The vector gravitational field at this point, due to m, is the acceleration of a free test particle there:

$$\mathbf{g} = -\nabla\Phi = -Gmr^{-3}\mathbf{r} \equiv -\Gamma_0\mathbf{r} \tag{2}$$

Finally, the gravity gradient tensor field due to m, is:

$$\boldsymbol{\Gamma} = \nabla\mathbf{g} = \Gamma_0\left(\frac{3\mathbf{r}\mathbf{r}^T}{r^2} - \mathbf{I}_3\right) \tag{3}$$

Outside the earth, the fields are closely approximated by these formulas. If the test mass is a spacecraft, in circular orbit about the earth at radius r, then the orbital angular velocity ω_o is given by:

$$\omega_o^2 = \Gamma_0 = \mu_e/r^3 \tag{4}$$

in which μ_e is the gravitational constant of the earth. The actual potential of the earth is complicated; but differs from (1) by only about 1 part in 1000 in low earth orbit, less at higher altitudes. The variations in turn are known to better than 1 part in 1000. Thus, if spacecraft attitude is actually inferred from gradiometer measurements, this error in knowledge of the field would lead to corresponding attitude determination errors on the order of 10^{-6} rad, almost surely not the worst error contribution. In any case, the intent of the study is to find the accuracy with which a gradiometer can measure attitude, *given that the field is known*; so the study neglects field knowledge errors.

On the other hand, neglect of the known deviation from sphericity (mainly oblateness) would lead to attitude errors on the order of 10^{-3} rad, usually unacceptable. However, our intent is to determine feasibility; so the form of the necessary oblateness correction is outside the scope. A real system would also have to deal with eccentric orbits; but as the orbit is not solved for, the observability of the attitude can't be seriously affected by eccentricity; and the spacecraft orbit is taken here as circular.

In general, coordinate systems are described by sets of right handed orthonormal base vectors $\mathbf{e}^x_\alpha$, where α = 1, 2, or 3 denotes the axis, and x indicates the system. 1st, the spacecraft system $\mathbf{e}^s$. This is the physical system in the spacecraft to which all the accelerometer input axes, and all other instruments, are aligned. For simplicity, it will be assumed that the origin of $\mathbf{e}^s$ is at the spacecraft center of mass. The term "spacecraft attitude" means the rotation that connects $\mathbf{e}^s$ to a trajectory system $\mathbf{e}^t$. $\mathbf{e}_1^t$ is defined as the local upward vertical, through the origin of $\mathbf{e}^s$, and $\mathbf{e}_3^t$ is parallel to the orbital angular momentum. $\mathbf{e}_2^t$ completes a right handed system, and is along the spacecraft velocity vector. $\mathbf{e}^t$ rotates uniformly at a rate ω_o about $\mathbf{e}_3^t$ relative to an inertial system that won't need to be identified further.

The connection between systems may be described by a matrix of direction cosines $\mathbf{A}$:

$$\mathbf{e}_\alpha^s = A_{\alpha\beta}\mathbf{e}_\beta^t \tag{5}$$

In this study, the spacecraft is assumed to be earth pointing; so $\mathbf{A}$ will be taken as a small rotation. It then can be expressed in terms of small yaw (ψ), roll (ϕ), and pitch (θ) angles; about $\mathbf{e}_1^t$, $\mathbf{e}_2^t$, and $\mathbf{e}_3^t$, respectively. In these terms, and to 1st order in the angles:

$$\mathbf{A} = \begin{bmatrix} 1 & \theta & -\phi \\ -\theta & 1 & \psi \\ \phi & -\psi & 1 \end{bmatrix} \tag{6}$$

The need for $\mathbf{e}^t$ is that the earth fields $\mathbf{g}$ and $\boldsymbol{\Gamma}$ are most conveniently expressed there:

$$\mathbf{g}^t = -\Gamma_0 r \mathbf{e}_1^t = \Gamma_0 r[-1,\ 0,\ 0]^T \tag{7}$$

$$\boldsymbol{\Gamma}^t = \Gamma_0 \,\mathrm{diag}\,[2,\ -1,\ -1] \tag{8}$$

and expressing these in $\mathbf{e}^s$, where the instruments reside:

$$\mathbf{g}^s = \mathbf{A}\mathbf{g}^t = \Gamma_0 r[-1,\ \theta,\ -\phi]^T \tag{9}$$

$$\boldsymbol{\Gamma}^s = \mathbf{A}\boldsymbol{\Gamma}^t\mathbf{A}^T = \Gamma_0 \begin{bmatrix} 2 & -3\theta & 3\phi \\ -3\theta & -1 & 0 \\ 3\phi & 0 & -1 \end{bmatrix} \tag{10}$$

again to 1st order in the angles. Note first, that while pitch and roll turn up in these expressions, yaw does not. Physically, this is because $\mathbf{r}$ is an axis of symmetry of the fields.

If we could measure either $\mathbf{g}$ or $\boldsymbol{\Gamma}$ in $\mathbf{e}^s$, we could infer both θ and ϕ. Alas, accelerometers don't measure gravitational acceleration at all, and gradiometers are strongly perturbed by angular velocities and accelerations (see below). What if dynamic effects could be removed? For example, if a spot measurement of Γ_{13}^s could somehow be made, the error in ϕ would be:

$$\delta\phi = \delta\Gamma/(3\Gamma_0) + 3\phi\delta r/r \tag{11}$$

Suppose an orbit altitude of 500 km. Then $r = 6.867 \times 10^6$ m, and $\Gamma_0 = 1231$ E. A gradient measurement accuracy of .01 E would then contribute 2.708×10^{-6} rad to $\delta\phi$. The analysis of $\delta\theta$ is similar, given a measurement of Γ_{12}^s. In each case, the 2nd contribution to the error comes from the uncertainty in the knowledge of r. Supposing $\delta r = 10$ m, and $\phi = 0.1$ rad, this contribution to $\delta\phi$ comes to 4.37×10^{-7} rad. Since satellite tracking usually determines r rather better; and attitude control is typically much tighter; the tracking contribution may be regarded as conservative, and won't be considered further. Thus, if spot measurements of the gradient could be made at the .01 E level, then roll and pitch determination at the microradian level would be possible. If this gradient measurement came from a pair of accelerometers, with an 0.5 m separation and independent errors, their required accuracy would be

$$\delta a = 0.5(10^{-11})/2^{1/2} = 3.536 \times 10^{-12}\ \mathrm{m/s}^2$$

within the capability of the best room temperature accelerometers today, operating in space.

3 Dynamic Attitude Estimation

If gradiometers actually measured the gradient, then a model would be something like $\mathbf{z} = \boldsymbol{\Gamma}$ plus noise, or a subset of its components. A least squares analysis would then yield the covariance of the errors in the estimate of the desired ϕ and θ, for each discrete sample $\mathbf{z}$. However, as any real gradiometer measurement $\mathbf{z}$ contains functions of ω and $\dot{\omega}$, least squares analysis won't suffice; and we have to resort to dynamic estimation. The plant equations consist of the Euler equations of more or less rigid body motion, plus kinematic equations relating ω to the attitude angular rates. Actually, as there is very little process noise (external torque variations), these equations add considerable strength to the estimates; thus turning a practical necessity into a virtue. These plant equations are developed and linearized below, a process noise model is spelled out, a filter is synthesized, and it's shown how the terminal covariance of the errors in the estimates may be determined. A few results are given.

A major variation from the earlier gradiometer dynamic estimation studies, [5] and [4], is that, instead of treating gradiometers as measuring the intrinsic tensor (see below), this study follows [8] in treating the instrument as an array of accelerometers. The measurement models consist of what each accelerometer should measure, plus noise. One advantage of this is that the measurement noises are now uncorrelated, avoiding the careful treatment needed in [5]. For simplicity, the spacecraft is supposed to be a box, with edges l_α aligned along the $\mathbf{e}^s_\alpha$. Supposing a uniform density ρ, the spacecraft mass is:

$$m = \rho l_1 l_2 l_3 \tag{12}$$

A typical density might be $\rho = 1000$ kg/m^3; and the principal moments of inertia are:

$$J_1 = m(l_2^2 + l_3^2)/12 \quad ; \quad J_2 = m(l_1^2 + l_3^2)/12 \quad ; \quad J_3 = m(l_1^2 + l_2^2)/12 \tag{13}$$

The orbit is assumed circular, at a radius r. Assuming an altitude of 500 km, $r = 6.867 \times 10^6$ m, $\omega_o = .0011095$ rad/s, and $\Gamma_0 = 1231$ E. Also, the spacecraft speed in orbit is $v_o = r\omega_o = 7614$ m/s.

In [4] it's shown that the Euler equations of rigid body motion, when modified to include an arbitrary bias momentum $\mathbf{h}_W$, can be written as:

$$\mathbf{J}\dot{\boldsymbol{\omega}} = (\mathbf{J}\boldsymbol{\omega} + \mathbf{h}_W) \times \boldsymbol{\omega} + \boldsymbol{\tau}_{gg} + \boldsymbol{\tau}_e \tag{14}$$

in which the external torque has been separated into the gravity gradient torque $\boldsymbol{\tau}_{gg}$ and the nongravitational torque $\boldsymbol{\tau}_e$, the latter mostly due to air drag. Note that the derivative is the rate of change as seen in $\mathbf{e}^s$. Control torques could be included in $\boldsymbol{\tau}_e$; but as they would then reappear in the filter structure equations, they cancel out in the covariance study. Unfortunately, this system is nonlinear in $\boldsymbol{\omega}$. Since we are analyzing a nominally earth pointing satellite, the nominal value of $\boldsymbol{\omega}$ is $\omega_o\mathbf{e}^t_3$. However, because of the body derivatives, a much simpler procedure is to define the variation $\boldsymbol{\epsilon}$ by:

$$\boldsymbol{\omega} = \omega_o \mathbf{e}^s_3 + \boldsymbol{\epsilon} \tag{15}$$

Another simplification comes by arguing that, in an earth pointing satellite, bias momentum, if any, is usually confined to the pitch axis:

$$\mathbf{h}_W = h\mathbf{e}^s_3 \tag{16}$$

Additional wheels for control aren't precluded; it's only required that their nominal momentum is zero. Substituting these relations into (14), and deleting quadratic terms in $\boldsymbol{\epsilon}$, yields

$$\mathbf{J}\dot{\boldsymbol{\epsilon}} = \omega_o(\mathbf{J}\boldsymbol{\epsilon}) \times \mathbf{e}^s_3 + \omega_o(\mathbf{J}\mathbf{e}^s_3) \times (\omega_o\mathbf{e}^s_3 + \boldsymbol{\epsilon}) + h\mathbf{e}^s_3 \times \boldsymbol{\epsilon} + \boldsymbol{\tau}_{gg} + \boldsymbol{\tau}_e \tag{17}$$

We also need $\boldsymbol{\tau}_{gg}$. The well known formula in $\mathbf{e}^t$ may be put in the form:

$$\boldsymbol{\tau}^t_{gg} = 3\Gamma_0 \mathbf{e}^t_1 \times (\mathbf{J}^t\mathbf{e}^t_1) \tag{18}$$

Since only $\mathbf{J}^s$ is readily available, and as what we really need is $\boldsymbol{\tau}^s_{gg}$, we need to work out

$$\boldsymbol{\tau}^s_{gg} = 3\Gamma_0\mathbf{A}[\mathbf{e}^t_1 \times (\mathbf{A}^T\mathbf{J}^s\mathbf{A}\mathbf{e}^t_1)] = 3\Gamma_0 \begin{bmatrix} -J_{12}\phi - J_{13}\theta \\ (J_{11} - J_{33})\phi + J_{23}\theta - J_{13} \\ (J_{11} - J_{22})\theta + J_{23}\phi + J_{12} \end{bmatrix} \tag{19}$$

Note that, while nothing depends on ψ, there is a yaw torque, arising from off diagonal components of $\mathbf{J}$. These also produce bias torques in roll and pitch. That's why, for earth pointing satellites, it's generally preferable to point some principal axis up. Moreover, by making this axis ($\mathbf{e}_1^s$) have the least J, $\boldsymbol{\tau}_{gg}$ is restoring. Here, where the main issue is observability, it's assumed that this condition is met, when

$$\mathbf{J}^s = \mathrm{diag}\,[J_1,\ J_2,\ J_3] \tag{20}$$

In the examples, it's further assumed that $J_1 < J_2 < J_3$, known to be the best configuration for gravity gradient stabilized satellites. With the principal axis assumption, the torque reduces to:

$$\boldsymbol{\tau}_{gg}^s = 3\Gamma_0[0,\ (J_1 - J_3)\phi,\ (J_1 - J_2)\theta]^T \tag{21}$$

On putting this into (17), and expressing it in standard form, the Euler equations become:

$$\begin{aligned} \dot{\epsilon}_1 &= k_1\epsilon_2 + J_1^{-1}\tau_{e1} \\ \dot{\epsilon}_2 &= k_2\epsilon_1 + k_3\phi + J_2^{-1}\tau_{e2} \\ \dot{\epsilon}_3 &= k_4\theta + J_3^{-1}\tau_{e3} \end{aligned} \tag{22}$$

in which the constants are defined as:

$$k_1 = [\omega_o(J_2 - J_3) - h]/J_1 \ ; \ k_2 = [\omega_o(J_3 - J_1) + h]/J_2 \ ; \ k_3 = 3\Gamma_o(J_1 - J_3)/J_2 \ ; \ k_4 = 3\Gamma_o(J_1 - J_2)/J_3 \tag{23}$$

If the gravity gradient and other external torques are neglected, then ϵ_1 and ϵ_2 decouple from ϵ_3 in (22), resulting in a harmonic oscillator with frequency ω_N given by:

$$\omega_N^2 = -k_1k_2 \tag{24}$$

This is the natural nutation frequency, arising mainly from the momentum bias h.

To complete the plant equations we must add the kinematical relations. With the same linearizing assumptions, these are easily shown to be (see for instance [4]):

$$\dot{\psi} = \epsilon_1 + \omega_o\phi \ ; \ \dot{\phi} = \epsilon_2 - \omega_o\psi \ ; \ \dot{\theta} = \epsilon_3 \tag{25}$$

We now have a linear system of plant equations of 6th order in $\boldsymbol{\epsilon}$, ψ, ϕ, and θ.

The random process appearing in the Euler equations (22) is the external non-gravitational torque $\boldsymbol{\tau}_e$. At 500 km, this is largely due to air drag; and the random component is largely from variations in air density ρ_a. For gradiometer studies, a flat earth barometric model was adopted in [4]:

$$\rho_a(r + \delta r) = \rho_a(r)e^{-\delta r/h_s} \tag{26}$$

where h_s is the density scale height. At 500 km, [9] lists $\rho_a = 1.905 \times 10^{-12}$ kg/m^3, $h_s = 83,000$ m, and a mean free path of 25,000 m. These numbers are, admittedly, quite shaky. In any case, the dynamic pressure then comes from the speed:

$$q = \rho_a v_o^2/2 \tag{27}$$

and with the above numbers, $q = 1.106 \times 10^{-4}$ N/m^2. Since the speed is along $\mathbf{e}_2^t$, and the spacecraft attitude is not far from nominal, the steady force from air drag is very nearly:

$$\mathbf{f}_e = -ql_1l_3C_D\mathbf{e}_2^s \tag{28}$$

Because the mean free path is much larger than the spacecraft, drag is essentially Newtonian, with a coefficient $C_D = 2$. However, since some inelastic, oblique, and diffuse scattering of air molecules is likely, this C_D may be high, and $C_D = 1.5$ is adopted. We should also consider radiation pressure. Corresponding to q is I_s/c, where $I_s = 1360$ w/m^2 the mean insolance outside the earth, and c is the speed of light. Thus, the mean "radiation dynamic pressure" is 4.54×10^{-6} N/m^2, well below q; and as the variations are much slower than for air drag, radiation pressure is ignored. [4] goes on to develop a statistical model. It supposes that ρ_a is actually the mean of a distribution, to which a random component is added:

$$\rho_r = \rho_a w_d(t) \tag{29}$$

$w_d(t)$ is a dimensionless, zero mean, random function of position and time. At satellite speed, the spatial variation dominates. Suppose that $w_d(t)$ has a standard deviation σ_w. Still, we need a power spectrum. Physically, we are looking at dynamic variations in density, with scale lengths of order h_s, plus the orbital frequency variation due to solar heating of the atmosphere. The latter, while reaching substantial amplitudes, is confined to such low frequencies as to have little effect on the attitude estimates, and is ignored. As for dynamic variations, we can imagine variability on all length scales, but petering out below distances of order h_s. This situation led to the development of the cubic power spectrum in [2]:

$$S(\omega) = \frac{\pi R(0)}{\omega_c}\left(1 - \frac{\omega}{2\omega_c}\right)^2\left(1 + \frac{\omega}{\omega_c}\right) \qquad (0 \le \omega \le 2\omega_c) \tag{30}$$

and zero otherwise. Suppose the autocorrelation of variations falls by half at a distance αh_s. The time to travel this distance is $\lambda = \alpha h_s/v_o$, and [2] shows that, for the cubic spectrum, we should choose:

$$\omega_{cw} = \frac{\pi}{2\lambda} = \frac{\pi v_o}{2\alpha h_s} \tag{31}$$

We must also pick $R_w(0)$ and α. The best information presently available to us is an analysis of CACTUS data in [10]. Accelerometer data over approximately 800 s intervals was analyzed at altitudes between 270 and 320 km. Density variations of $\sim$ 4%, peak to peak were typical; rising sometimes to $\sim$ 15%, during severe magnetic disturbances. The corresponding σ_w values are .014 and .05. A reasonable balance between these values would be $\sim$ .02; but, allowing for a bit greater variability at higher altitudes, we have taken $\sigma_w = .025$. Then, as these time series meet the oversampling conditions discussed in the Appendix, $R_w(0) = \sigma_w^2 = 6.25 \times 10^{-4}$. As for α, [10] doesn't show a power spectrum, but does give representative time series of a normal and a disturbed interval; and states that the apparent wavelengths concentrate in the range of 700 to 1500 km. Examination of the time series suggests that $R(\tau)$ falls to 0.5 at $\tau \sim 50$ s. Translating to our altitude, the corresponding distance is 381 km, when $\alpha = 4.6$. Since for a sinusoid, $R(\tau)$ falls by half at 1/3 of a wavelength, these numbers are at least consistent. Again, to allow for a bit more variability at 500 km, we have taken $\alpha = 4$, leading to $\omega_{cw} = 0.03606$ rad/s.

It remains to convert this to torque. The overall drag force is very nearly:

$$\mathbf{f}_e = -k_f[1 + w_d(t)]\mathbf{e}_2^s \tag{32}$$

where

$$k_f \equiv \rho_a v_o^2 l_1 l_3 C_D/2 \tag{33}$$

Supposing a center of pressure at a location $\mathbf{r}_{cp}$ in the spacecraft, the torque due this is:

$$\boldsymbol{\tau}_e = \mathbf{r}_{cp} \times \mathbf{f}_e = k_f[1 + w_d(t)][r_{cp3},\ 0,\ -r_{cp1}]^T \tag{34}$$

Note that there a deterministic bias force and torque, which must be treated correctly in the filter. Also, while our box structure has no torque along $\mathbf{e}_2^s$, an actual spacecraft would likely have a small propeller torque on this axis. To allow for this below, a component r_{cp2} replaces the zero in (34).

4 Measurement Model

In [5] and [3], the instrument was modeled as measuring elements of the "intrinsic" tensor:

$$\mathbf{T} = \boldsymbol{\Gamma} + \omega^2\mathbf{I}_3 - \omega\omega^T + \boldsymbol{\varepsilon}\dot{\omega} \tag{35}$$

where $\boldsymbol{\varepsilon}$ is the 3–index permutation symbol. The quadratic ω terms are centrifugal effects. Because the instrument is fixed in $\mathbf{e}^s$, there is no coriolis. Here, the instrument is dissolved into its component accelerometers, partly to avoid the noise correlations required in [5] and [3], but mainly to prepare for later studies. The gradiometer is taken as an array of 3 axis accelerometers, with input axes aligned along the $\mathbf{e}_\alpha^s$. For entering symmetrical arrays, it's convenient to identify a "center" of the instrument $\mathbf{r}_c$, relative to the origin of $\mathbf{e}^s$. Then, the ith accelerometer will have a position $\mathbf{r}_{ai}$, relative to the center. Thus, its location relative to the center of mass is:

$$\mathbf{r}_i = \mathbf{r}_c + \mathbf{r}_{ai} \tag{36}$$

For a perfectly circular orbit, the center of mass is subject to $-\omega_o^2 r\mathbf{e}_1^t$. As for rotation effects, $\mathbf{e}^s$ is rotating at a rate ω, relative to an inertial frame $\mathbf{e}^n$. So, purely due to rotation, the inertial velocity of the ith accelerometer is (the superscripts indicate the frame in which the derivative is observed):

$$\dot{\mathbf{r}}_i = \frac{d^n}{dt}\mathbf{r}_i = \frac{d^s}{dt}\mathbf{r}_i + \boldsymbol{\omega} \times \mathbf{r}_i = \boldsymbol{\omega} \times \mathbf{r}_i \tag{37}$$

the latter because $\mathbf{r}_i$ is invariant in $\mathbf{e}^s$. Going to the next derivative

$$\ddot{\mathbf{r}}_i = \frac{d^n}{dt}\dot{\mathbf{r}}_i = \dot{\boldsymbol{\omega}} \times \mathbf{r}_i + \boldsymbol{\omega} \times \frac{d^n}{dt}\mathbf{r}_i = \dot{\boldsymbol{\omega}} \times \mathbf{r}_i + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}_i) \tag{38}$$

Note that $\dot{\boldsymbol{\omega}}$ is the same, whether viewed from $\mathbf{e}^n$ or $\mathbf{e}^s$. Finally, on including the external non-gravitational acceleration $\mathbf{a}_e$, the ith accelerometer is subject to:

$$\mathbf{a}_i = -\omega_o^2 r\mathbf{e}_1^t + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}_i) + \dot{\boldsymbol{\omega}} \times \mathbf{r}_i + \mathbf{a}_e \tag{39}$$

On the other hand, the gravitational acceleration of the ith accelerometer is $\mathbf{g}^t$ plus the correction at $\mathbf{r}_i$ due to the gradient. From (7) and (4), this comes to:

$$\mathbf{g}_i = -\omega_o^2 r\mathbf{e}_1^t + \boldsymbol{\Gamma}\mathbf{r}_i \tag{40}$$

Actual accelerometers measure only non-gravitational acceleration; i.e., the difference between inertial and gravitational acceleration. These are identical in free fall, when an accelerometer measures zero. Conversely, an accelerometer on a table on earth measures the acceleration imposed by the table that keeps the instrument from falling through the floor. Thus, the ith accelerometer model is:

$$\mathbf{z}_i = \mathbf{a}_i - \mathbf{g}_i + \mathbf{v}_i \tag{41}$$

where $\mathbf{v}_i$ is the noise in the 3 measurements. On substituting from above this becomes:

$$\mathbf{z}_i = \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}_i) + \dot{\boldsymbol{\omega}} \times \mathbf{r}_i - \boldsymbol{\Gamma}\mathbf{r}_i + \mathbf{a}_e + \mathbf{v}_i \tag{42}$$

Note that the acceleration of the center of mass has dropped out. The next step is to linearize this using (15). On neglecting the quadratic terms, and recalling that $\dot{\boldsymbol{\omega}}$ is the same in $\mathbf{e}^n$ and $\mathbf{e}^s$, we get:

$$\mathbf{z}_i = \omega_o(\omega_o\mathbf{e}_3^s + \boldsymbol{\epsilon}) \times (\mathbf{e}_3^s \times \mathbf{r}_i) + \omega_o\mathbf{e}_3^s \times (\boldsymbol{\epsilon} \times \mathbf{r}_i) + \dot{\boldsymbol{\epsilon}} \times \mathbf{r}_i - \boldsymbol{\Gamma}\mathbf{r}_i + \mathbf{a}_e + \mathbf{v}_i \tag{43}$$

We'll work this out term by term, in the form of matrices of constants times the state variables, plus whatever is left over. Starting on the left:

$$\mathbf{e}_3^s \times (\mathbf{e}_3^s \times \mathbf{r}_i) = r_{i3}\mathbf{e}_3^s - \mathbf{r}_i = -[r_{i1},\ r_{i2},\ 0]^T \tag{44}$$

$$\boldsymbol{\epsilon} \times (\mathbf{e}_3^s \times \mathbf{r}_i) = (\mathbf{r}_i \cdot \boldsymbol{\epsilon})\mathbf{e}_3^s - \epsilon_3\mathbf{r}_i = \begin{bmatrix} 0 & 0 & -r_{i1} \\ 0 & 0 & -r_{i2} \\ r_{i1} & r_{i2} & 0 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} \tag{45}$$

$$\mathbf{e}_3^s \times (\boldsymbol{\epsilon} \times \mathbf{r}_i) = r_{i3}\boldsymbol{\epsilon} - \epsilon_3\mathbf{r}_i = \begin{bmatrix} r_{i3} & 0 & -r_{i1} \\ 0 & r_{i3} & -r_{i2} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} \tag{46}$$

The $\dot{\boldsymbol{\epsilon}}$ term can't be expressed directly in the state variables; however, from (22), there follows:

$$\dot{\boldsymbol{\epsilon}} \times \mathbf{r}_i = \begin{bmatrix} k_2r_{i3} & 0 & k_3r_{i3} & -k_4r_{i2} \\ 0 & -k_1r_{i3} & 0 & k_4r_{i1} \\ -k_2r_{i1} & k_1r_{i2} & -k_3r_{i1} & 0 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \phi \\ \theta \end{bmatrix} + \begin{bmatrix} J_2^{-1}r_{i3}\tau_{e2} - J_3^{-1}r_{i2}\tau_{e3} \\ J_3^{-1}r_{i1}\tau_{e3} - J_1^{-1}r_{i3}\tau_{e1} \\ J_1^{-1}r_{i2}\tau_{e1} - J_2^{-1}r_{i1}\tau_{e2} \end{bmatrix} \tag{47}$$

The $\boldsymbol{\Gamma}$ term comes directly from (10):

$$\boldsymbol{\Gamma}\mathbf{r}_i = 3\Gamma_o \begin{bmatrix} r_{i3} & -r_{i2} \\ 0 & -r_{i1} \\ r_{i1} & 0 \end{bmatrix} \begin{bmatrix} \phi \\ \theta \end{bmatrix} + \Gamma_o \begin{bmatrix} 2r_{i1} \\ -r_{i2} \\ -r_{i3} \end{bmatrix} \tag{48}$$

and on combining all these, and substituting from the process noise model:

$$\mathbf{z}_i = \begin{bmatrix} (k_3+\omega_o)r_{i3} & 0 & -2\omega_o r_{i1} & 0 & (k_3-3\Gamma_o)r_{i3} & (3\Gamma_o-k_4)r_{i2} \\ 0 & (\omega_o-k_1)r_{i3} & -2\omega_o r_{i2} & 0 & 0 & (k_4+3\Gamma_o)r_{i1} \\ (\omega_o-k_2)r_{i1} & (k_1+\omega_o)r_{i1} & 0 & 0 & -(k_3+3\Gamma_o)r_{i1} & 0 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \psi \\ \phi \\ \theta \end{bmatrix}$$

$$+\ \Gamma_o \begin{bmatrix} -3r_{i1} \\ 0 \\ r_{i3} \end{bmatrix} + k_f \begin{bmatrix} J_2^{-1}r_{i3}r_{cp2} + J_3^{-1}r_{i2}r_{cp1} \\ -J_3^{-1}r_{i1}r_{cp1} - J_1^{-1}r_{i3}r_{cp3} - m^{-1} \\ J_1^{-1}r_{i2}r_{cp3} - J_2^{-1}r_{i1}r_{cp2} \end{bmatrix} [1+w_d(t)] + \mathbf{v}_i \tag{49}$$

This completes the description of the accelerometers. There is 1 such 3 vector for each accelerometer. The noise depends critically on instrument design; but as we are interested only in feasibility, no particular instrument is used. Since a power spectrum is needed even for a generic instrument, a cubic spectrum similar to (30) is assumed. The Appendix shows how the average power $R_v(0)$, and the break frequency ω_{cv}, are determined from the rms acceleration error and the averaging time τ of the measurement.

5 Filter Structure

The 1st step in calculating the terminal covariance in a dynamic estimation problem is to determine the structure of the filter. This starts with identifying the set of state variables that appear in the plant and measurement equations. From (22) and (25), it's clear that we should choose:

$$\mathbf{x} = [\epsilon_1,\ \epsilon_2,\ \epsilon_3,\ \psi,\ \phi,\ \theta]^T \tag{50}$$

Following [7], it's conventional to consolidate the plant equations in standard linearized form:

$$\dot{\mathbf{x}} = \mathbf{F}\mathbf{x} + \mathbf{G}(\mathbf{u}) + \mathbf{B}\mathbf{w} \tag{51}$$

Here, $\mathbf{F}$ is the plant matrix, $\mathbf{u}$ is a vector of controls, $\mathbf{G}(\mathbf{u})$, a possibly nonlinear vector function, distributes the controls, $\mathbf{w}$ is a vector of independent process noises, and $\mathbf{B}$ is the process noise state distribution matrix. The matrices are readily identified. From (22) and (25), we find:

$$\mathbf{F} = \begin{bmatrix} 0 & k_1 & 0 & 0 & 0 & 0 \\ k_2 & 0 & 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & k_4 \\ 1 & 0 & 0 & 0 & \omega_o & 0 \\ 0 & 1 & 0 & -\omega_o & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{52}$$

As for the control and process noise terms, it's convenient to separate the deterministic process noise bias from the random components, and combine them with the actual controls, if any, to produce the $\mathbf{G}(\mathbf{u})$ used here. Since these terms will eventually cancel out in the analysis below, the actual controls have no effect on filter performance, and there is no need to spell out $\mathbf{G}(\mathbf{u})$. Finally, by identifying $\mathbf{w}$ with $w_d(t)$ in (34), and including propeller torque, we have:

$$\mathbf{B} = k_f[r_{cp3}/J_1,\ r_{cp2}/J_2,\ -r_{cp1}/J_3,\ 0,\ 0,\ 0]^T \tag{53}$$

Turning now to the measurement model, the direct appearance of the process noise in each of the accelerometer measurements requires a modification of the usual standard model:

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{Y}\mathbf{v} + \mathbf{U}\mathbf{w} + \mathbf{z}_B \tag{54}$$

Here, $\mathbf{H}$ is the measurement partials matrix, developed above. From (49), this is:

$$\mathbf{H}_i = \begin{bmatrix} (k_3+\omega_o)r_{i3} & 0 & -2\omega_o r_{i1} & 0 & (k_3-3\Gamma_o)r_{i3} & (3\Gamma_o-k_4)r_{i2} \\ 0 & (\omega_o-k_1)r_{i3} & -2\omega_o r_{i2} & 0 & 0 & (k_4+3\Gamma_o)r_{i1} \\ (\omega_o-k_2)r_{i1} & (k_1+\omega_o)r_{i2} & 0 & 0 & -(k_3+3\Gamma_o)r_{i1} & 0 \end{bmatrix} \tag{55}$$

and the complete measurement partials matrix is:

$$\mathbf{H} = [\mathbf{H}_1^T,\ \mathbf{H}_2^T,\ \cdots]^T \tag{56}$$

For example, if there are 7 vector accelerometers, $\mathbf{H}$ will be a 21×6 matrix.

For measurement noise, it's assumed that each axis of each accelerometer has separate independent noise. Thus, $\mathbf{v}(t)$ has one element for each element of $\mathbf{z}$, and $\mathbf{Y}$ is just an identity. A more elaborate model may be found in [1]; so $\mathbf{Y}$ is retained in what follows. The spectral properties of $\mathbf{v}(t)$ were developed above. As for the process noise term, having established that $\mathbf{w}$ is $w_d(t)$, $\mathbf{U}$ comes immediately from (49):

$$\mathbf{U}_i = k_f \begin{bmatrix} J_2^{-1} r_{i3} r_{cp2} + J_3^{-1} r_{i2} r_{cp1} \\ -J_3^{-1} r_{i1} r_{cp1} - J_1^{-1} r_{i3} r_{cp3} - m^{-1} \\ J_1^{-1} r_{i2} r_{cp3} - J_2^{-1} r_{i1} r_{cp2} \end{bmatrix} \tag{57}$$

The overall $\mathbf{U}$ is a column vector with 3 such elements for each accelerometer. The remaining terms in (49) constitute the bias $\mathbf{z}_B$. As it doesn't affect the covariance analysis below, it's not spelled out.

An observer based on these models starts with an estimate $\hat{\mathbf{x}}$ of the state $\mathbf{x}$. This is caused to follow the deterministic parts of the plant equations (51), corrected by feeding back the residuals, i.e., the actual measurements $\mathbf{z}$ minus the measurement model (54). The filter structure then takes the form:

$$\dot{\hat{\mathbf{x}}} = \mathbf{F}\hat{\mathbf{x}} + \mathbf{G}(\mathbf{u}) + \mathbf{K}(\mathbf{z} - \mathbf{H}\hat{\mathbf{x}} - \mathbf{z}_B) \tag{58}$$

Note that this structure assumes that the control and bias terms are known, and available to the filter. The issue buried here is that $\mathbf{G}(\mathbf{u})$ is accurately modeled, and that the biases have been accurately determined by some sort of in flight calibration. Pursuing these points is beyond our scope.

6 Terminal Covariance

The performance of a dynamic filter is generally examined by determining the statistics of the error in the estimate, defined by:

$$\boldsymbol{\xi} \equiv \hat{\mathbf{x}} - \mathbf{x} \tag{59}$$

The evolution of $\boldsymbol{\xi}$ comes from subtracting (51) from (58):

$$\dot{\boldsymbol{\xi}} = \mathbf{Z}\boldsymbol{\xi} + \mathbf{K}\mathbf{Y}\mathbf{v}(t) - \mathbf{W}\mathbf{w}(t) \tag{60}$$

where the observer system matrix and the process noise effect matrix are defined by:

$$\mathbf{Z} \equiv \mathbf{F} - \mathbf{K}\mathbf{H} \quad ; \quad \mathbf{W} \equiv \mathbf{B} - \mathbf{K}\mathbf{U} \tag{61}$$

There's lots to learn from (60). 1st, $\mathbf{x}$, $\hat{\mathbf{x}}$, and all the control and bias terms have disappeared. Thus, the quality of the estimate doesn't depend on the controls, even if they fail to stabilize the plant — the "Separation Theorem" in the controls business. 2nd, filter stability requires $\mathbf{Z}$ to be stable; i.e., all its eigenvalues are in the left half plane, a standard requirement in any negative feedback system. Filter theory puts this differently: if a $\mathbf{K}$ can be found such that $\mathbf{Z}$ is stable, then the state $\mathbf{x}$ is said to be observable by the measurements $\mathbf{z}$. 3rd, the diagonal elements and the eigenvalues of $\mathbf{Z}$ have the dimensions of inverse time; and filter settling time is essentially given by the inverse of its least negative eigenvalue. This is used below to insure that the "optimal" filter has a reasonable settling time. Finally, since the noises are unbiased, so is $\boldsymbol{\xi}(t)$.

Various measures have been proposed to study the quality of the estimate. Here, and generally in the references, attention has centered on the covariance of the error:

$$\mathbf{P}_\xi(t) \equiv E[\boldsymbol{\xi}(t)\boldsymbol{\xi}^T(t)] \tag{62}$$

where E is the expectation operator. The idea that, in a stationary system, $\mathbf{P}_\xi(t)$ would have a terminal or asymptotic value, has been around a long time, but finding it could be quite tedious, if the settling time was long. About 4 years ago, William McEneaney, in unpublished notes, showed that this terminal value

$\mathbf{P}_\xi$ could be calculated directly from the structural information and the noise statistics. On generalizing to arbitrary power spectra, his ideas led to [6] and [7].

The present problem differs from [7] primarily by including process noise in the measurement model. Also, [7] dealt with the autocovariances of all the noises, and it has since been found much easier to work with power spectra directly. Since none of this appears in print, the algorithm for calculating $\mathbf{P}_\xi$ is derived here. To begin, it may be supposed that the filter has been running for all past time; when the initial conditions have settled out. Then (60) is solved for "now" in this form:

$$\boldsymbol{\xi}(0) = \int_0^\infty e^{\mathbf{Z}\mu}[\mathbf{KYv}(\mu) - \mathbf{Bw}(\mu)]d\mu \tag{63}$$

where the dummy variable μ may be interpreted as past time. Strictly, the noise terms should be $v(-\mu)$ and $w(-\mu)$; but, as only the statistical properties of $\boldsymbol{\xi}$ matter, it makes no difference. An apparently graver problem is $e^{\mathbf{Z}\mu}$ — the dimensions of $\mathbf{Z}\mu$ depend on those of $\mathbf{x}$, thus calling into question the validity of the formal expansion. However, from (60), the dimensions of the vector $t\mathbf{Zx}$ are just those of $\mathbf{x}$. Thus, all terms of the form $\mu^i\mathbf{Z}^i\mathbf{x}$ have the same dimensions, and if the exponential is merely viewed as shorthand for the formal expansion, there are no dimensional difficulties.

The terminal covariance may now be found by substituting this into (62):

$$\mathbf{P}_\xi = \int_0^\infty \int_0^\infty e^{\mathbf{Z}\mu}\{\mathbf{KY}E[\mathbf{v}(\mu)\mathbf{v}^T(\nu)]\mathbf{Y}^T\mathbf{K}^T + \mathbf{W}E[\mathbf{w}(\mu)\mathbf{w}^T(\nu)]\mathbf{W}^T\}e^{\mathbf{Z}^T\nu}d\mu d\nu \tag{64}$$

This supposes that the expectation and integration operators may be commuted, and uses the assumption that $\mathbf{w}$ and $\mathbf{v}$ are independent and free of bias. On recognizing the autocorrelations of the noises, this is:

$$\mathbf{P}_\xi = \int_0^\infty \int_0^\infty e^{\mathbf{Z}\mu}[\mathbf{KYR}_v(\mu-\nu)\mathbf{Y}^T\mathbf{K}^T + \mathbf{WR}_w(\mu-\nu)\mathbf{W}^T]e^{\mathbf{Z}^T\nu}d\mu d\nu \tag{65}$$

Well, autocorrelations and power spectra are Fourier transforms of each other. Using the one sided spectra of [6], these relations for any noise component are:

$$R(\tau) = \frac{1}{\pi}\int_0^\infty S(\omega)\mathrm{c}(\tau\omega)d\omega \quad ; \quad S(\omega) = 2\int_0^\infty R(\eta)\mathrm{c}(\omega\tau)d\tau \tag{66}$$

After using the former in (65), and interchanging the order of integrations, there follows:

$$\mathbf{P}_\xi = \frac{1}{\pi}\int_0^\infty \int_0^\infty \int_0^\infty e^{\mathbf{Z}\mu}\mathbf{Q}(\omega)e^{\mathbf{Z}^T\nu}\mathrm{c}[\omega(\mu-\nu)]d\mu d\nu d\omega \tag{67}$$

in which:

$$\mathbf{Q}(\omega) \equiv \mathbf{KYS}_v(\omega)\mathbf{Y}^T\mathbf{K}^T + \mathbf{WS}_w(\omega)\mathbf{W}^T \tag{68}$$

Considerable progress can now be made by a change of coordinates:

$$\theta \equiv \mu + \nu \quad ; \quad \eta \equiv \mu - \nu \tag{69}$$

the double integration region is now the quadrant surrounding the $+\theta$ axis, so

$$\begin{aligned} \mathbf{P}_\xi \;=\; & \frac{1}{2\pi}\int_0^\infty \left[\int_{-\infty}^0 e^{\mathbf{Z}\eta/2}\int_{-\eta}^\infty e^{\mathbf{Z}\theta/2}\mathbf{Q}(\omega)e^{\mathbf{Z}^T\theta/2}d\theta e^{-\mathbf{Z}^T\eta/2}\mathrm{c}(\omega\eta)d\eta \right. \\ & + \left. \int_0^\infty e^{\mathbf{Z}\eta/2}\int_\eta^\infty e^{\mathbf{Z}\theta/2}\mathbf{Q}(\omega)e^{\mathbf{Z}^T\theta/2}d\theta e^{-\mathbf{Z}^T\eta/2}\mathrm{c}(\omega\eta)d\eta\right] d\omega \end{aligned} \tag{70}$$

Now, it's not hard to establish that

$$\int e^{\mathbf{Z}\theta/2}\mathbf{Q}(\omega)e^{\mathbf{Z}^T\theta/2}d\theta = 2e^{\mathbf{Z}\theta/2}\mathbf{D}(\omega)e^{\mathbf{Z}^T\theta/2} + \text{constant} \tag{71}$$

where $\mathbf{D}(\omega)$ satisfies the Lyapunov equation:

$$\mathbf{Z}\mathbf{D}(\omega) + \mathbf{D}(\omega)\mathbf{Z}^T = \mathbf{Q}(\omega) \tag{72}$$

Putting this into (70), setting $\eta \to -\eta$ in the 1st integral, and evaluating at the required limits, a considerable simplification results:

$$\mathbf{P}_\xi = -\frac{1}{\pi}\int_0^\infty \left[\mathbf{D}(\omega)\int_0^\infty e^{\mathbf{Z}^T\eta}c(\omega\eta)d\eta + \int_0^\infty e^{\mathbf{Z}\eta}c(\omega\eta)d\eta\,\mathbf{D}(\omega)\right] d\omega \tag{73}$$

when another analytic integral has surfaced:

$$\int_0^\infty e^{\mathbf{Z}\eta}c(\omega\eta)d\eta = -(\mathbf{Z} + \omega^2\mathbf{Z}^{-1})^{-1} \tag{74}$$

leading finally to:

$$\mathbf{P}_\xi = \frac{1}{\pi}\int_0^\infty [\mathbf{N}(\omega) + \mathbf{N}^T(\omega)]d\omega \quad ; \quad \mathbf{N}(\omega) \equiv (\mathbf{Z} + \omega^2\mathbf{Z}^{-1})^{-1}\mathbf{D}(\omega) \tag{75}$$

It may be noted that this analysis would break down in several places but for $\mathbf{Z}$ being stable. Once again, especially in (74), the dimensions may look flaky. However, letting u_i represent the dimensions of x_i, it is readily shown from the differential equations that the expressions $Z_{ij}t$, Z_{ij}/ω, and ωZ_{ij}^{-1} all have the dimensions u_i/u_j. By extension, the ijth element of (74) has the dimensions tu_i/u_j. This work has established the forward procedure. For a given $\mathbf{K}$, $\mathbf{Z}$ and $\mathbf{W}$ are computed from (61) and (61). A set of ω values is chosen to cover the region where any of the noise spectra are nonzero, with reasonable density. $\mathbf{Q}(\omega)$ is then determined over this set from (68). Each $\mathbf{Q}(\omega)$ yields a corresponding $\mathbf{D}(\omega)$ by solution of the Lyapunov equation (72), and a corresponding $\mathbf{N}(\omega)$ from (75). $\mathbf{P}_\xi$ is then found by integrating (75).

7 Optimal Feedback Gains

Having found how to compute $\mathbf{P}_\xi$ from $\mathbf{K}$, we still need to find the $\mathbf{K}$ that yields optimal filter performance, whatever that means. While $\mathbf{P}_\xi$ certainly contains the necessary information, in this 6th order problem there are 21 independent matrix elements; so some sort of scalar measure of $\mathbf{P}_\xi$ is needed. The software used here is based on a performance index q, constructed from the weighted trace of $\mathbf{P}_\xi$:

$$q = P_{\xi\alpha\alpha}/C_\alpha^2 \tag{76}$$

In this technique, known as "Bryson weighting", each C_i is the "level of concern" for the error ξ_i. For example, if x_i were a position, the level of concern might be $C_i = 1$ m. $C_i = 10$ m would show less concern, and cause the optimization to put less weight on the variance of ξ_i. Note that the Bryson technique has the virtue that q is the sum of dimensionless terms — it doesn't add apples and oranges.

A further concern can be added to the performance index — filter settling time. If the $\mathbf{K}$ that minimizes (76) leads to a $\mathbf{Z}$ with a small (though negative) eigenvalue, then we may see from (60) that the settling time of the filter will be long, perhaps excessively so. To avoid such a problem, a term may be added to (76) penalizing the filter settling time. To see how to do this, consider the behavior of the filter evolution equation (60). If λ_α symbolizes the eigenvalues of $\mathbf{Z}$, and $\sigma_\alpha \equiv \Re(\lambda_\alpha)$, then the filter response to initial conditions or perturbations may be regarded as a set of n exponentially decaying modes, with individual settling times $-1/\sigma_\alpha$. Since all n modes decay simultaneously, the overall settling time is:

$$t_s = \max_\alpha[-\Re(\lambda_\alpha)]^{-1} = -\left[\max_\alpha \Re(\lambda_\alpha)\right]^{-1} \tag{77}$$

Now suppose we introduce a concern level C_t in seconds for the settling time t_s. Then the overall performance index may be taken as:

$$q = (P_{\xi\alpha\alpha}/C_\alpha^2) + t_s/C_t \tag{78}$$

The added term serves another function. The stability boundary for $\mathbf{Z}$ is that all $\sigma_\alpha < 0$. Thus, as some $\sigma_\alpha \to 0$ from the left, $t_s \to \infty$. So, adding the t_s term erects a barrier against $\mathbf{Z}$ going unstable. If we have picked the concerns C_i and C_t, and have a $\mathbf{K}$, such that $\mathbf{Z}$ is stable, then $\mathbf{P}_\xi$ may be found as detailed above, and q computed. Next, each element of $\mathbf{K}$ is varied, to get a δq. Taken together, these constitute a ∇q, relative to the elements of $\mathbf{K}$. A minimum q is then found by searching along $-\nabla q$. This whole process is iterated until q bottoms out. The final $\mathbf{K}$ is the "optimal" feedback gain, and the final $\mathbf{P}_\xi$ and t_s comprise the filter performance at that gain. However, this result could just be local.

$C_{4,5,6}$ weight the variances in the attitude angular errors. Since our sponsor feels that 5×10^{-5} rad is a reasonable goal for θ and ϕ, these are the adopted concern levels, with .001 rad for ψ. As for $C_{1,2,3}$, an uncontrolled gravity stabilized satellite might sway by .05 rad at frequencies of order $2\omega_o$. Thus, the actual rates ϵ would be $\sim 10^{-4}$ rad/s. If we needed to know them to, say, 1%, our level of concern would be 10^{-6} rad/s; and this is taken as the concern level. However, an unusually stringent rate jitter requirement, would shrink the rate concern levels.

There is one serious loose end — the starting $\mathbf{K}$ must yield a stable $\mathbf{Z}$. The method we use is based on Kalman theory. Suppose each noise component $S(\omega)$ is replaced by a flat bounded spectrum with the same average power $R(0)$, and with cutoff frequency Ω where $S(\omega)$ vanishes for good. This level is $S = \pi R(0)/\Omega$. The white noise "equivalent" to $S(\omega)$ has level S out to infinity. Replacing all the noise components with these "equivalents" causes $\mathbf{Q}$, and thus $\mathbf{D}$ to be independent of ω. This allows $\mathbf{P}_\xi$ to be integrated analytically, leading to $\mathbf{P}_\xi = -\mathbf{D}$, when there is a clean connection between $\mathbf{K}$ and $\mathbf{P}_\xi$. On reorganizing with the help of (61) and (61), so as to make the dependence on $\mathbf{K}$ explicit, we have:

$$\mathbf{KHP}_\xi + \mathbf{P}_\xi\mathbf{H}^T\mathbf{K}^T - \mathbf{FP}_\xi - \mathbf{P}_\xi\mathbf{F}^T = \mathbf{KMK}^T - \mathbf{KV}^T - \mathbf{VK}^T + \mathbf{BS}_w\mathbf{B}^T \tag{79}$$

where:

$$\mathbf{M} \equiv \mathbf{YS}_v\mathbf{Y}^T + \mathbf{US}_w\mathbf{U}^T \quad ; \quad \mathbf{V} \equiv \mathbf{BS}_w\mathbf{U}^T \tag{80}$$

Since an optimum $\mathbf{P}_\xi$ is necessarily stationary relative to variations in $\mathbf{K}$, (79) may be expressed in components, and differentiated relative to each $K_{\mu\nu}$, leading to this stationarity condition for $\mathbf{P}_\xi$:

$$\mathbf{KM} = \mathbf{P}_\xi\mathbf{H}^T + \mathbf{V} \tag{81}$$

While this can't be used directly to eliminate either $\mathbf{K}$ or $\mathbf{P}_\xi$ from (79), we need only assume that some noise contaminates every measurement component to insure that $\mathbf{M}$ is non–singular. Thus:

$$\mathbf{K} = (\mathbf{P}_\xi\mathbf{H}^T + \mathbf{V})\mathbf{M}^{-1} \tag{82}$$

which, except for the $\mathbf{V}$ term, is a staple of Kalman theory. When this is substituted back into (79), an equally well known algebraic Riccati equation emerges:

$$\mathbf{A} + \mathbf{XP}_\xi + \mathbf{P}_\xi\mathbf{X}^T = \mathbf{P}_\xi\mathbf{H}^T\mathbf{M}^{-1}\mathbf{HP}_\xi \equiv \mathbf{P}_\xi\mathbf{LP}_\xi \tag{83}$$

where:

$$\mathbf{A} \equiv \mathbf{B}(\mathbf{S}_w - \mathbf{S}_w\mathbf{U}^T\mathbf{M}^{-1}\mathbf{US}_w)\mathbf{B}^T \quad ; \quad \mathbf{X} \equiv \mathbf{F} - \mathbf{VM}^{-1}\mathbf{H} \tag{84}$$

All this reduces to Kalman theory when the measurements don't depend on $\mathbf{w}(t)$; i.e., $\mathbf{U} = \mathbf{V} = \mathbf{0}$. In the software, (83) is solved for $\mathbf{P}_\xi$, and $\mathbf{K}$ is computed from (82). While this $\mathbf{K}$ is far from optimal for the real power spectra, it does guarantee a stable $\mathbf{Z}$ to start the iteration. A potential difficulty is that the Riccati equation has many solutions; but it's known that at most 1 yields $\mathbf{P} > \mathbf{0}$.

This is quite a large optimization. For example, if the gradiometer is composed of 4 vector accelerometers, $\mathbf{K}$ has 72 elements, all of which must be determined. Such problems are touchy, and the difficulties are aggravated by poor conditioning in $\mathbf{P}_\xi$ or $\mathbf{Z}$. Some sort of scaling is usually applied to alleviate this. Here, a natural scaling already exists — the Bryson concern levels. On the hypothesis that the variance $P_{\xi ii}$ is on the same order of magnitude as C_i^2, consider scaling the state variables and time:

$$x_i' = x_i/C_i \quad ; \quad t' = t/C_t \tag{85}$$

which are non-dimensional. Recall that, in the convention adopted in this paper, summation is only over lower case *greek* indices. The covariance of the scaled variables is then:

$$P'_{\xi ij} = E(x'_i x'_j) = P_{\xi ij}/(C_i C_j) \tag{86}$$

The real virtue of such a scaling is that the eigenvalues of $\mathbf{P}'_\xi$ should be much closer together than those of $\mathbf{P}_\xi$, with a corresponding improvement in the condition number. To carry out this scaling, (85) is substituted into (51), leading to:

$$\dot{\mathbf{x}}' = \mathbf{F}'\mathbf{x}' + \mathbf{G}'(\mathbf{u}) + \mathbf{B}'\mathbf{w} \tag{87}$$

in which

$$F'_{ij} = F_{ij}C_t C_j/C_i \quad ; \quad B'_{ij} = B_{ij}C_t/C_i \quad ; \quad G'_i = G_i C_t/C_i \tag{88}$$

It's not hard to show that the scaling makes all these arrays dimensionless. While it's not necessary to scale the measurements, in the model we must set

$$\mathbf{H}\mathbf{x} = \mathbf{H}'\mathbf{x}' \tag{89}$$

from which

$$H'_{ij} = C_j H_{ij} \tag{90}$$

The filter structure then becomes:

$$\dot{\hat{\mathbf{x}}}' = \mathbf{F}'\hat{\mathbf{x}}' + \mathbf{G}'(\mathbf{u}) + \mathbf{K}'(\mathbf{z} - \mathbf{H}'\hat{\mathbf{x}}' - \mathbf{z}_B) \tag{91}$$

in which the derivatives are with respect to t', and

$$K'_{ij} = K_{ij}C_t/C_i \tag{92}$$

and the error in the estimate

$$\xi'_i = \hat{x}'_i - x'_i = \xi_i/C_i \tag{93}$$

evolves as

$$\dot{\boldsymbol{\xi}}' = \mathbf{Z}'\boldsymbol{\xi}' + \mathbf{K}'\mathbf{Y}\mathbf{v}(t) - \mathbf{W}'\mathbf{w}(t) \tag{94}$$

where

$$\mathbf{W}' \equiv \mathbf{B}' - \mathbf{K}'\mathbf{U} \quad ; \quad \mathbf{Z}' \equiv \mathbf{F}' - \mathbf{K}'\mathbf{H}' \tag{95}$$

In components, these matrices are related to the unscaled versions by:

$$W'_{ij} = W_{ij}C_t/C_i \quad ; \quad Z'_{ij} = Z_{ij}C_t C_j/C_i \tag{96}$$

Note that the matrices $\mathbf{Y}$ and $\mathbf{U}$, and thus $\mathbf{M}$ aren't affected by scaling. From the determinant relation for eigenvalues, it's not hard to show that those of $\mathbf{Z}'$ obey

$$\lambda'_\alpha = C_t\lambda_\alpha \Rightarrow \sigma'_\alpha = C_t\sigma_\alpha \Rightarrow t'_s = t_s/C_t \tag{97}$$

On substituting these scaling relations into (78), q becomes rather simple:

$$q = \mathrm{Tr}(\mathbf{P}'_\xi) + t'_s \tag{98}$$

The modified iteration starts by forming $\mathbf{B}'$ and $\mathbf{F}'$. Then, the transformed algebraic Riccati equation is

$$\mathbf{A}' + \mathbf{X}'\mathbf{P}'_\xi + \mathbf{P}'_\xi\mathbf{X}'^T = \mathbf{P}'_\xi\mathbf{L}'\mathbf{P}'_\xi \tag{99}$$

in which $\mathbf{A}'$, $\mathbf{X}'$, and $\mathbf{L}'$ are computed as above, except that $\mathbf{F}$, $\mathbf{B}$, and $\mathbf{H}$ are replaced by their primed equivalents. Note that $\mathbf{V} \rightarrow \mathbf{V}'$, but no scaling of $\mathbf{M}$ is required. Solving this leads to a starting value $\mathbf{P}'_\xi$ for the main iteration. Applying the scaling everywhere, the iteration becomes:

$$\mathbf{Q}'(\omega) \equiv \mathbf{K}'\mathbf{Y}\mathbf{S}_v(\omega)\mathbf{Y}^T\mathbf{K}'^T + \mathbf{W}'\mathbf{S}_w(\omega)\mathbf{W}'^T \tag{100}$$

The Lyapunov equation is then:

$$\mathbf{Z}'\mathbf{D}'(\omega) + \mathbf{D}'(\omega)\mathbf{Z}'^T = \mathbf{Q}'(\omega) \tag{101}$$

whose solution leads to $\mathbf{N}'$ and $\mathbf{P}'_\xi$. Finally, when q has settled, yielding the terminal $\mathbf{K}'$ and $\mathbf{P}'_\xi$, the unscaled values are

$$P_{\xi ij} = C_i C_j P'_{\xi ij} \quad ; \quad K_{ij} = C_i K'_{ij}/C_t \tag{102}$$

8 Results

The calculation of the terminal covariance for a given set of input data requires the exercise of several programs in sequence, all more or less interactive. The programs are all written in APL, and implemented on a 486DX 33 Mhz computer. A typical run requires several hours, almost entirely for the minimization of q, but including all the interactive input and output routines. All the results cited here are based on the numbers in the text. The spacecraft dimensions are 2.0, 0.7, and 0.5 m; with a mass of 140 kg. A momentum bias of 10 N–m–s is added, yielding a natural nutation frequency of 0.48458 rad/s. From the air data in the text, ω_{cwd} = .036046 rad/s. For numerical integration, 63 points were used in the ω vector; but a couple of runs were repeated with more points, to insure the accuracy. The instrument consists of 4 accelerometers at the corners of a regular tetrahedron, whose circumscribed sphere has a radius of 0.25 m. The noise levels ranged from 2×10^{-10} to 10^{-8} m/s^2. The averaging time was 1 s, for ω_{cv} = 62.832 rad/s. The C_i are as in the text; C_t = 10 s.

In all cases, q was dominated by t_s; although this dominance wanes at higher measurement noise levels. Our interpretation is that ψ is observable only through roll–yaw coupling in the kinematic equations, at a natural frequency ω_o, and thus causes long settling times. As the noise increases, so must $\mathbf{P}_\xi$; and t_s rises to maintain the concern balance. Presumably, a filter simulation would show that roll and pitch would settle much more quickly. This behavior is seen in the following table:

$10^9\times$ rms error – m/s^2	0.2	0.5	1	2	5	10
σ yaw – μrad	56.3	11.9	34.5	51.0	22.8	82.1
σ roll – μrad	3.04	3.33	7.06	32.0	20.1	72.5
σ pitch – μrad	4.18	2.92	3.74	16.7	34.5	138
t_s – s	223	315	336	670	695	1378

The progression to higher noise seems rather erratic. We believe that this is due in part to the t_s dominance, but much more to the $\lambda(\mathbf{Z})$. In the 2nd and 3rd runs, t_s comes from complex twins. The 1st and 4th run produced triplets, 1 of which was real; while the last 2 runs yielded quadruplets, composed of 2 complex pairs. In most cases, coalescence signaled that further iteration is unproductive. In all cases there were dramatic improvements from the Riccati equation starting $\mathbf{K}$ to the final value. Clearly, there is a great deal of room for further research; and many more runs are planned, varying other parameters.

The authors would like to acknowledge the considerable assistance of Prof. Penina Axelrad of the Univ. of Colorado. Most of the work was performed under a contract to Analytical Engineering from the Univ. of Colorado, in turn supported by a Grant from Goddard Space Flight Center.

Appendix — Averaged Measurement Noise

The instruments studied here are modeled as measuring the acceleration of their case, plus noise. In practice, they generally average the analog output for some time interval τ, and deliver a digital result after each interval. The study considers only analog instruments, and thus takes $\tau = 0$. On the other hand, the instrument manufacturers often characterize their devices as delivering "samples" (really averages) every τ seconds, or alternatively, at a sample rate of $1/\tau$ Hz. The noise associated with these averages is then specified by a standard deviation σ. This appendix deals with relating this type of specification to the parameters of the assumed cubic power spectrum. This situation was examined in [11], where it was found that for an arbitrary noise power spectrum $S(\omega)$, the variance of the averages is:

$$\sigma^2 = \frac{2}{\pi\tau^2}\int_0^\infty S(\omega)[1 - \mathrm{c}(\tau\omega)]\frac{d\omega}{\omega^2} \tag{103}$$

Assuming the cubic spectrum (30) for the analog noise, the variance can be put in the form:

$$\sigma^2 = R(0) f_s(\tau\omega_c) \tag{104}$$

where, in terms of the sine integral function:

$$f_s(x) = \frac{2}{x}\mathrm{Si}(2x) + \frac{\mathrm{s}x}{x^3}\left(\frac{\mathrm{s}x}{x} + \mathrm{c}x\right) - \frac{2}{x^2}(1 + \mathrm{s}^2 x) \tag{105}$$

in which:

$$\mathrm{Si}(y) = \int_0^y \frac{\mathrm{s}z}{z} dz = y - \frac{y^3}{3 \cdot 3!} + \frac{y^5}{5 \cdot 5!} - \cdots \tag{106}$$

The function looks ghastly for $x \ll 1$; but it actually behaves quite well:

$$f_s(x) = 1 - \frac{x^2}{9} + O(x^4) \tag{107}$$

This is the oversampling limit; i.e., if a time series is very frequently measured, but is long enough to cover many cycles of the highest noise frequency, then $R(0)$ is the variance of the samples, and the distinction between sample and average disappears. Actually, this limit holds for any $S(\omega)$. The other limit, $x \gg 1$ is also clean: $\mathrm{Si}(x) \to \pi/2$ and $f_s(x) \to \pi/x$. Overall, $f_s(x)$ is a monotonic decreasing function, whose behavior can be seen from the table:

x	0	0.1	0.2	0.5	1	2	3
$f_s(x)$	1	0.99956	0.99823	0.98901	0.9574	0.84917	0.71822
x	5	10	20	50	100	200	500
$f_s(x)$	0.50907	0.28422	0.14958	.061632	.031116	.015633	.006271

When σ was measured by the manufacturer, the repetition frequency $1/\tau$ was probably chosen about an order of magnitude below the half power frequency $\omega_c/(2\pi)$. Adapting this reasoning, we can pick:

$$\omega_c = 20\pi/\tau \tag{108}$$

so that $\tau\omega_c = 20\pi = 62.832$ rad; and $R(0) = .0492401\sigma^2$. This assumed structure has been used to determine the measurement noise power spectrum in the study.

References

[1] D Sonnabend, "A Simple GRADIO Accelerometer Model", JPL EM 314–495, 1–24–91.

[2] D Sonnabend, "Cubic Power Spectra", JPL EM 314–507, 6–18–91.

[3] AM San Martín, "SGGM In Flight Calibration", JPL IOM 343–90–601, 11–3–90.

[4] D Sonnabend, "Realistic Gradiometer Dynamics", Proc. IUGG Meeting, Vienna Austria, 8–91; also JPL EM 314–508, 7–26–91.

[5] D Sonnabend & WM McEneaney, "Gravity Gradient Measurements", Proc. IEEE Meeting on Decision and Control, Bierman Memorial Symposium, Austin TX, 12–88.

[6] D Sonnabend, "Who Needs Markov?", JPL EM 314–478, 6–8–90.

[7] D Sonnabend, "Terminal Covariance", JPL EM 314–481, 6–11–90.

[8] D Sonnabend, "Locating a Free Mass", JPL EM 314–465, 9–17–90.

[9] CW Allen, "Astrophysical Quantities", 2nd Ed., Athlone Press, 1963.

[10] F Barlier & C Berger, "Rapid density variations — Statistical analysis; Applications to gradiometry", CIGAR CERGA 140/CMTR/s, 6-12-88.

[11] D Sonnabend, "Time Averaged Noise", JPL EM 314–494, 12–12–90.